

Genome-wide Association Studies in Alzheimer's Disease: A Review

Giuseppe Tosto · Christiane Reitz

Published online: 17 August 2013

© Springer Science+Business Media New York 2013

Abstract Over the past decade, research aiming to disentangle the genetic underpinnings of late-onset Alzheimer's disease has mostly focused on the identification of common variants through genome-wide association studies. The identification of several new susceptibility genes through these efforts has reinforced the importance of amyloid precursor protein and tau metabolism in the cause of the disease and has implicated immune response, inflammation, lipid metabolism, endocytosis/intracellular trafficking, and cell migration in the cause of the disease. Ongoing and future large-scale genome-wide association studies, translational studies, and next-generation whole genome or whole exome sequencing efforts, hold the promise to map the specific causative variants in these genes, to identify several additional risk variants, including rare and structural variants, and to identify novel targets for genetic testing, prevention, and treatment.

Keywords Alzheimer's disease · Genetics · Gene · Variation · Polymorphism · Genome-wide association study · Sequencing

Introduction

Late-onset Alzheimer's disease (LOAD) typically begins with the onset of symptoms after the age of 60 years

and evolves slowly from mildly impaired memory to severe cognitive loss. At death, the most frequent pathological manifestations in the brain include extracellular β -amyloid protein (A β) in diffuse and neuritic plaques and intracellular deposits of hyperphosphorylated tau protein, a microtubule assembly protein, in the form of neurofibrillary tangles. Widespread loss of both neurons and synapses also occurs [1].

An estimated 4.5 million Americans have LOAD. The annual incidence of LOAD increases from 1 % at the age of 60–70 years to 10–30 % at 85 years and older [2]. As the US population ages, it is expected that the number of LOAD cases will increase to 16 million 20 million by 2050, with one in 45 Americans affected [3, 4]. A critical barrier to lessening the impact of this disease is the limited development of drugs to prevent or treat LOAD, which is mostly attributable to incomplete characterization of the basic underlying pathologic mechanisms. Determining which genes and gene networks contribute to LOAD risk would reveal basic pathogenic mechanisms, highlighting key proteins and pathways for drug development ("druggable targets"), and inform the development of genetic testing methods for identifying those at greatest risk of LOAD when preventive measures become available.

In recent years, the genetic analysis of LOAD has focused on identification of common variants through genome-wide association studies (GWAS) and has identified several novel susceptibility genes implementing specific pathways in the disease. This article reviews these studies, discusses their potentials and limitations, and provides suggestions for future research.

Data Source and Study Selection

The primary sources of the studies addressed in this review were full-text articles and abstracts published in English in the PubMed database between 2010 and February 2013. The keywords used for searching PubMed were

This article is part of the Topical Collection on *Dementia*

C. Reitz

Taub Institute for Research on Alzheimer's Disease and the Aging Brain, College of Physicians and Surgeons, Columbia University, New York, NY, USA

G. Tosto · C. Reitz

Department of Neurology, College of Physicians and Surgeons, Columbia University, New York, NY, USA

C. Reitz (✉)

Gertrude H. Sergievsky Center, College of Physicians and Surgeons, Columbia University, 630 West 168th Street, New York, NY 10032, USA
e-mail: cr2101@columbia.edu

“dementia,” “Alzheimer’s disease,” “gene,” “genetics,” “epigenetics,” “endophenotype,” and “genome-wide association study.” The abstracts retrieved were read to identify studies addressing the topics included in this review. We also performed a manual search of references cited in published articles. The studies were read in their entirety to assess their appropriateness for inclusion in this article.

Genetic Epidemiology of LOAD

A family history of dementia is one of the most important risk factors for LOAD [5, 6]. Families multiply affected by LOAD are at increased risk of dementia, but the distribution of secondary cases is not consistent with Mendelian inheritance. LOAD is more frequent among monozygotic twins than dizygotic twins [7–9], and first-degree relatives of patients with LOAD have approximately twice the expected lifetime risk of developing the disease. Heritabilities of 58–79 % for LOAD indicate that in spite of progress made in identifying the underpinnings of the disease, a substantial fraction of LOAD is attributable to unknown genetic factors.

Apolipoprotein E Region

For more than a decade, only one genetic risk factor, the *APOE* $\epsilon 4$ allele, located on chromosome band 19q13, was an unequivocally established “susceptibility” gene in non-Hispanic Whites of European ancestry. Apolipoprotein E (ApoE) is a lipid-binding protein and is expressed in humans as three common isoforms coded for by three alleles, $\epsilon 2$, $\epsilon 3$, and $\epsilon 4$. A single *APOE* $\epsilon 4$ allele is associated with a twofold to threefold increased risk; having two copies is associated with a fivefold or more increased risk [10]. In addition, each inherited *APOE* $\epsilon 4$ allele lowers the age at onset by 6–7 years [11–18]. *APOE* $\epsilon 4$ is also associated with lower cognitive performance, in particular the memory domain, is associated with mild cognitive impairment [19–22], and is associated with progression from mild cognitive impairment to dementia [19–29]. Although the population attributable risk of *APOE* $\epsilon 4$ is estimated at 20–50 % [30], the presence of $\epsilon 4$ is neither necessary nor sufficient for development of the disease [31]. In ethnic groups other than non-Hispanic Whites, the association between *APOE* and LOAD was largely inconsistent across studies.

Findings from GWAS

At the beginning of the century, thousands of candidate-gene-based association studies aiming to identify additional susceptibility loci were performed, but only one gene, the sortilin-related receptor (*SORL1*) [32], which is implicated in intracellular trafficking of amyloid precursor protein (APP), could be

consistently replicated in independent datasets and implicated in the disease. The main reasons for these inconsistencies between studies are sample heterogeneity with differences in linkage disequilibrium (LD) patterns and allele frequencies, and small sample sizes, leading to limited power to detect small or moderate effect sizes. In the past 5 years, technological advances in high-throughput genome-wide arrays have allowed the hypothesis-free simultaneous examination of thousands to millions of polymorphisms across the genome, and large collaborative efforts capitalizing on this technology have significantly advanced knowledge of the genetic underpinnings of LOAD and the pathways involved by identifying several novel risk loci.

The major GWAS contributing to this gained knowledge are summarized in Table 1. Most were performed in non-Hispanic Whites of European ancestry. The first set of studies identified *CLU*, *PICALM*, *CR1*, and *BIN1* as susceptibility loci [33–35]. Clusterin (Clu), also known as apolipoprotein J, is a lipoprotein highly expressed in both the periphery and the brain [36]. Like ApoE, it is involved in lipid transport [37]. Clu is also hypothesized to act as an extracellular chaperone that influences A β aggregation and receptor-mediated A β clearance by endocytosis [38]. Unlike for *APOE*, there are no known coding variants that account for the observed genetic association to *CLU*, suggesting that genetic variation in expression levels may be responsible for the altered risk of LOAD [38]. *BIN1* (amphiphysin II) is a member of the Bin1/amphiphysin/RVS167 (BAR) family of genes that are involved in diverse cellular processes, including actin dynamics, membrane trafficking, and clathrin-mediated endocytosis [39] which affect APP processing and A β production or A β clearance from the brain. Phosphatidylinositol-binding clathrin assembly protein (*PICALM*) is also involved in clathrin-mediated endocytosis and recruits clathrin and adaptor protein complex 2 to sites of vesicle assembly [40]. *CR1* is a cell-surface receptor that is part of the complement system. It has binding sites for complement factors C3b and C4b and is involved in clearing immune complexes containing these two proteins. Since A β oligomers can bind C3b, *CR1* may participate in the clearance of A β . *CR1* may also play a role in neuroinflammation, which is a prominent feature in Alzheimer’s disease [41]. Interestingly, Clu may play a role in this process as an inhibitor [42]. In summary, this first set of GWAS identified loci mainly clustering in four pathways, namely, immune response, APP processing, lipid metabolism, and endocytosis/intracellular trafficking.

The second set of large GWAS identified additional susceptibility genes (*CD33*, *MS4A4A/MS4A4E/MS4A6E* cluster, *ABCA7*, *CD2AP*, and *EPHA1*) [43•, 44•]. In line with the pathways identified by the first set of GWAS, all of these five loci are likely involved in the immune system, whereas *ABCA7* is in addition involved in lipid metabolism and APP processing (Table 2). The *CD33* gene encodes a protein that is

Table 1 Major Alzheimer's disease (AD) genome-wide association studies (GWAS) performed

Study	Ethnic group	Sample size	Genes identified outside <i>APOE</i> region
Lambert et al. [34]	Caucasian	Stage 1: 2,032 AD cases; 5,328 controls Stage 2: 3,978 AD cases; 3,297 controls	<i>CLU, CR1</i>
Harold et al. [33]	Caucasian	Stage 1: 3,941 AD cases; 7,848 controls Stage 2: 2,023 AD cases; 2,340 controls	<i>CLU, PICALM</i>
Seshadri et al. [35]	Caucasian	Stage 1: 3,006 AD cases; 4,642 controls Stage 2: 2,032 AD cases; 5,328 controls Stage 3: 3,333 AD cases; 6,995 controls	<i>BIN1, XOC3L2/BLOC1S3/MARK4, CLU, PICALM</i>
Naj et al. [44••]	Caucasian	Stage 1: 8,309 AD cases; 7,366 controls Stage 2: 3,531 AD cases; 3,565 controls	<i>MS4A4A, CD2AP, CD33, EPHA1, CR1, CLU, BIN1, PICALM</i>
Hollingworth et al. [43••]	Caucasian	Stage 1: 6,688 AD cases; 13,685 controls Stage 2: 4,896 AD cases; 4,903 controls Stage 3: 8,286 AD cases; 21,258 controls	<i>ABCA7, MS4A6A/MS4A4E, EPHA1, CD33, CD2AP</i>
Lee et al. [58]	Caribbean Hispanic	549 AD cases; 544 controls	<i>CLU, PICALM, BIN1, CUGBP2</i> , loci on 2p25.1; 3q25.2; 7p21.1; 10q23.1
Reitz et al. [59••]	African American	1,968 AD cases; 3,928 controls	<i>ABCA7</i> , intergenic locus on 5q35.2

a member of a family of cell-surface immune receptors that bind extracellular sialylated glycans and signal via a cytoplasmic domain called the immunoreceptor tyrosine inhibitory motif [45, 46]. CD33 has primarily been studied in the peripheral immune system, where it is expressed on myeloid progenitors and monocytes and also in the brain. In the periphery, CD33 appears to inhibit proliferation of myeloid cells [47]. The *MS4A4A/MS4A4E/MS4A6E* locus is part of a cluster of 15 *MS4A* genes on chromosome 11 and encodes proteins with multiple membrane-spanning domains that were initially identified by their homology to CD20, a B-lymphocyte cell-surface molecule. Little is known about the function of *MS4A4A* gene products; however, like CD33, *MS4A4A* is expressed on myeloid cells and monocytes and likely has an immune-related function. *EPHA1* encodes a member of the ephrin family of cell-surface receptors which interact with ephrin ligands on adjacent cells to modulate cell adhesion, migration, and axon guidance and synapse formation and plasticity. Although there is a substantial body of research

on the function of ephrin receptors in general, little is known about the *EPHA1* gene product. Like other ephrin receptors, it regulates cell morphology and motility [48] and early work implicated this receptor in regulating vascular morphogenesis and angiogenesis [49]. *EPHA1* knockout in mice results in abnormal tail and reproductive tract development [50], but no effects on the brain. Consistent with this notion, in mice, expression is restricted to epithelial tissue. In humans, *EPHA1* is expressed by CD4⁺ T lymphocytes [51], monocytes [52], intestinal epithelium, and colon. Combined with the lack of evidence for brain expression, this may suggest that, like CD33, CR1, and *MS4A4/MS4A6E*, the role of the *EPHA1* gene product in Alzheimer's disease may be mediated through the immune system. The CD2-associated protein gene (*CD2AP*) encodes a scaffolding protein that binds directly to actin [53], nephrin, and other proteins involved in cytoskeletal organization. In the immune system, CD2AP is required for synapse formation [54] in a process that involves clathrin-dependent actin polymerization. *ABCA7* is an integral transmembrane ATP-binding cassette (ABC) transporter belonging to the ABC family of proteins that mediate the biogenesis of high-density lipoprotein with cellular lipid and helical apolipoproteins [55]. It binds apolipoprotein A-I and functions in apolipoprotein-mediated phospholipid and cholesterol efflux from cells [56]. In addition, *ABCA7* affects the transport of other important proteins, including APP [56], through the cell membrane and is involved in host defense through effects on phagocytosis by macrophages of apoptotic cells [55].

In these large-scale GWAS performed in non-Hispanic Whites of European ancestry, the most strongly associated single-nucleotide polymorphisms (SNPs) at each locus other than *APOE* demonstrated population attributable fractions between 1.0 and 8.0 %, with effect sizes ranging from an odds

Table 2 Major pathways identified by GWAS

Pathway	Gene
Amyloid pathway	<i>APOE, SORL1, CLU, CR1, PICALM, BIN1, ABCA7</i>
Immune system/inflammation	<i>CLU, CR1, EPHA1, ABCA7, MS4A4A/MS4A6E, CD33, CD2AP</i>
Lipid transport and metabolism	<i>APOE, CLU, ABCA7</i>
Synaptic cell functioning/endocytosis	<i>CLU, PICALM, BIN1, EPHA1, MS4A4A/MS4A6E, CD33, CD2AP</i>

ratio of 1.16 to an odds ratio of 1.20, i.e., much smaller than for *APOE* [57]. In the largest GWAS performed to date in Caribbean Hispanics [58], associations in *CLU*, *PICALM*, and *BIN1* were replicated and several additional loci on 2p25.1, 3q25.2, 7p21.1, and 10q23.1—which could be replicated in an independent cohort of non-Hispanic Whites of European ancestry from the National Institute on Aging Late-Onset Alzheimer's Disease Family Study (NIA-LOAD)—were observed. Finally, in the largest GWAS of African Americans performed, Reitz et al. [59•] identified *ABCA7* as a major susceptibility locus in this ethnic group. Interestingly, in contrast to all GWAS loci identified in Caucasians, in African Americans the *ABCA7* locus had an effect size as strong as that of *APOE* $\epsilon 4$ (i.e., a 70–80 % increase in risk compared with a 10–20 % increase in risk through the GWAS loci observed in Whites). Although this finding may represent a winner's curse (i.e., inflation of the estimated effect in a discovery set in relation to follow-up studies) and needs to be confirmed by independent studies in African Americans and functional methods, it may have major implications for developing targets for genetic testing, prevention, and treatment in this ethnic group if proven true. In addition, this study confirmed *APOE* as a susceptibility gene in this ethnic group, evidence for which prior to this study had been inconsistent across studies, and also replicated *CRI*, *BIN1*, *EPHA1*, and *CD33*.

Discussion

The recent GWAS for LOAD using large numbers of cases and controls identified several novel susceptibility loci that are biologically plausible, cluster in specific pathways, and have significantly advanced the understanding of the pathogenic mechanism underlying the disease. Common to all novel loci in non-Hispanic Whites of European ancestry is the modest effect size with odds ratio ranging from 1.1 to 1.5 leaving the *APOE* $\epsilon 4$ allele by far the strongest risk factor. In contrast, in the largest GWAS performed to date in African Americans, the *ABCA7* locus was observed to have an effect size similar to that of *APOE* (70–80 % increase in risk). The population attributable risk of each of the non-*APOE* loci is estimated to be 1–8 %. However, this estimate will change with elucidation of the number, allele frequencies, and risk effects of the true functional variants at each locus, and the detection of additional common and rare risk variants and patterns of epistasis.

Replication in independent datasets—if possible across different ethnic groups—and functional validation of the loci identified by GWAS is crucial for several reasons. First, GWAS are not designed to identify the specific causative variants, but rather are designed screen the genome, capitalizing on the LD between genotyped SNPs and the potentially

causative variants [60]. As LD can extend over large intervals, the true genetic effectors may be located considerably far away from the SNP showing the disease association, limiting the ability to detect true associations from GWAS. The development of high-throughput genotyping arrays, which have increased the number of genotyped markers to several millions, has decreased this problem to some extent, but not entirely, depending on the LD pattern in the region. Second, signals selected on the basis of statistical significance thresholds in underpowered settings are often subject to the winner's curse (bias away from the null in the estimated effect of a newly identified allele on disease) [61–63], and replication can help produce a more accurate, unbiased estimate of the genetic effect of a locus. Third, the probability that an observed association truly exists depends on the power to detect the association, which in turn is a function of minor allele frequency, effect size, sample size, and the observed *p* value. The distribution of effect sizes of true associations in complex diseases is unknown, but it is likely that most of the large effects in LOAD GWAS have been identified, whereas most of the smaller effects remain to be discovered. The significance threshold needed to preserve the genome-wide type I error rate in studies of individuals with European ancestry is estimated at 5×10^{-7} – 1×10^{-8} [64, 65]. This threshold is even lower in ethnic groups with greater genetic diversity such as Hispanics, Africans, and African Americans and, consequently, most individual GWAS do not have enough power to distinguish false positives from false negatives. Finally, replication in a population with different environmental or genetic backgrounds may—if assessed in a population with a lower extent of LD such as Africans—help narrow down the location of the causative variant. In addition, it allows one to determine the generalizability of the observed association. However, when the aim is to replicate an observed association, it has to be kept in mind that there are several reasons for the observation of no association, including differences in allele frequencies or LD patterns across populations, or allelic or locus heterogeneity.

There are several additional approaches that can address some of these issues inherent to GWAS. Reclassifying sample subjects into more homogeneous subgroups, for example, based on endophenotypes, can reduce phenotypic heterogeneity and increase the power to detect true associations.

Gene-based association studies, which consider association between a trait and all markers within a gene rather than each marker individually, can be more powerful than traditional individual-SNP-based GWAS. For example, if a gene contains more than one causative variant, then several SNPs within that gene might show marginal levels of significance that are often indistinguishable from random noise in the initial GWAS results. If the effects of all SNPs in a gene are combined into a test statistic and correction is made for LD, the gene-based test might be able to detect these effects. Similarly, genome-

wide haplotype-based association studies can characterize loci not detected by univariate analyses. Such gene-based or haplotype-based analyses led to the discovery of *NARS2*, *FRMD6*, and *FRMD4A* as susceptibility loci [66, 67], the latter of which is immediately adjacent to *GAB2*. Identification and examination of regions with runs of homozygosity (i.e., excess burden of homozygous markers) can help identify recessive causative genes. Evidence is accumulating that a substantial part of the missing genetic variability could be due to epistatic effects or gene–environment interactions. Thus, exploration of gene–gene and gene–environment interactions can identify novel variants not detected by individual testing of SNPs. However, such studies require large samples sizes and/or large effect sizes to achieve adequate power.

Although the latest GWAS arrays include dense SNP maps of several million SNPs with minor allele frequency down to 1 % and novel functional exonic variants that were identified through sequencing of thousands of exomes, they are limited in their ability to detect associations with variants not tagged by the genotyped SNPs. In addition, they are limited in their ability to identify structural variants or rare variants with minor allele frequency of less than 1 %. However, both rare and structural variants are increasingly recognized as being implicated in complex disease [68]. In fact, two recent studies that performed genome sequencing followed by imputation of identified variants in independent datasets implicated the triggering receptor expressed on myeloid cells 2 gene (*TREM2*) in Alzheimer's disease by identifying a causative rare missense mutation (rs75932628) resulting in an R47H substitution affecting the gene's anti-inflammatory function [69, 70]. Additional sequencing studies identified rare causative variants in the nicastrin gene (*NCSTN*) encoding an obligatory component of the γ -secretase complex involved in splicing of APP [71] as well as *CLU* [72]. Although individual rare variants may have an effect size large enough to cause disease, the accumulation of several rare variants each with small or modest effect sizes may cross the susceptibility threshold. Ongoing and future large-scale next-generation whole exome or whole genome sequencing techniques will fill this gap and further provide the means to identify the specific causative variants in the genes/regions identified by GWAS. Although appropriate algorithms for the statistical and bioinformatic analysis of sequencing data, in particular for whole genome sequencing data and whole exome or whole genome sequencing data derived from families, still need to be developed and implemented, the recent identification of rare variants in *CLU*, *NCSTN*, and *TREM2* in LOAD that also cluster in amyloid processing and immune-response/inflammation pathways and were missed by the GWAS but identified by sequencing studies clearly belie the common disease–common variant hypothesis and prove the necessity of approaches with the ability to detect rare variants [69, 70, 71, 72]. Once causative variants are identified, functional studies can assess the pathogenic effects

of the variants and characterize the molecular pathways in which they are involved or with which they interact, further implicating the gene in the disease and potentially providing targets for effective intervention.

Conclusions and Future Directions

Over the past 10 years, studies capitalizing on high-throughput genome technologies have significantly advanced knowledge of the genetic underpinnings of LOAD. GWAS have identified several susceptibility genes, and sequencing studies have identified specific causative variants in these genes, but have also provided invaluable evidence for an involvement of rare variants in this complex disease, overturning the common disease–common variant hypothesis that had long defined the genetic research of complex diseases. Ongoing and future large-scale next-generation sequencing approaches (both hypothesis-driven and hypothesis-free) are likely to disentangle a significant part of the missing heritability of LOAD, and have the potential to identify targets for genetic testing, prevention, and treatment.

Acknowledgments Christiane Reitz was supported by a Paul B. Beeson Career Development Award (K23AG034550) and has received grant support from the National Institutes of Health/National Institute on Aging. Giuseppe Tosto was supported by a Defense Medical Research and Development Program (DMRDP) grant (W81XWH).

Compliance with Ethics Guidelines

Conflict of Interest Giuseppe Tosto and Christiane Reitz declare that they have no conflict of interest.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Papers of particular interest, published recently, have been highlighted as:

- Of importance
- Of major importance

1. Duyckaerts C, Delatour B, Potier MC. Classification and basic pathology of Alzheimer disease. *Acta Neuropathol*. 2009;118:5–36.
2. Mayeux R. Epidemiology of neurodegeneration. *Annu Rev Neurosci*. 2003;26:81–104.
3. Brookmeyer R, Gray S, Kawas C. Projections of Alzheimer's disease in the United States and the public health impact of delaying disease onset. *Am J Public Health*. 1998;88:1337–42.
4. Sloane PD, Zimmerman S, Suchindran C, Reed P, Wang L, Boustani M. The public health impact of Alzheimer's disease, 2000–2050: potential implication of treatment advances. *Annu Rev Public Health*. 2002;23:213–31.

5. Lautenschlager NT, Cupples LA, Rao VS, Auerbach SA, Becker R, Burke J, et al. Risk of dementia among relatives of Alzheimer's disease patients in the MIRAGE study: what is in store for the oldest old? *Neurology*. 1996;46:641–50.
6. Mayeux R, Sano M, Chen J, Tatemichi T, Stern Y. Risk of dementia in first-degree relatives of patients with Alzheimer's disease and related disorders. *Arch Neurol*. 1991;48:269–73.
7. Bergem AL, Lannfelt L. Apolipoprotein E type $\epsilon 4$ allele, heritability and age at onset in twins with Alzheimer disease and vascular dementia. *Clin Genet*. 1997;52:408–13.
8. Gatz M, Pedersen NL, Berg S, Johansson B, Johansson K, Mortimer JA, et al. Heritability for Alzheimer's disease: the study of dementia in Swedish twins. *J Gerontol A Biol Sci Med Sci*. 1997;52:M117–25.
9. Gatz M, Reynolds CA, Fratiglioni L, Johansson B, Mortimer JA, Berg S, et al. Role of genes and environments for explaining Alzheimer disease. *Arch Gen Psychiatry*. 2006;63:168–74.
10. Kuusisto J, Koivisto K, Kervinen K, Mykkanen L, Helkala EL, Vanhanen M, et al. Association of apolipoprotein E phenotypes with late onset Alzheimer's disease: population based study. *BMJ*. 1994;309:636–8.
11. Breitner JC, Wyse BW, Anthony JC, Welsh-Bohmer KA, Steffens DC, Norton MC, et al. APOE- $\epsilon 4$ count predicts age when prevalence of AD increases, then declines: the Cache County Study. *Neurology*. 1999;53:321–31.
12. Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science*. 1993;261:921–3.
13. Gomez-Isla T, West HL, Rebeck GW, Harr SD, Growdon JH, Locascio JJ, et al. Clinical and pathological correlates of apolipoprotein E $\epsilon 4$ in Alzheimer's disease. *Ann Neurol*. 1996;39:62–70.
14. Holmes C, Levy R, McLoughlin DM, Powell JF, Lovestone S. Apolipoprotein E: non-cognitive symptoms and cognitive decline in late onset Alzheimer's disease. *J Neurol Neurosurg Psychiatry*. 1996;61:580–3.
15. Kurz A, Altland K, Lautenschlager N, Zimmer R, Busch R, Gerundt I, et al. Apolipoprotein E type 4 allele and Alzheimer's disease: effect on age at onset and relative risk in different age groups. *J Neurol*. 1996;243:452–6.
16. Murman DL, Foster NL, Kilgore SP, McDonagh CA, Fink JK. Apolipoprotein E and Alzheimer's disease: strength of association is related to age at onset. *Dementia*. 1996;7:251–5.
17. Poirier J, Davignon J, Bouthillier D, Kogan S, Bertrand P, Gauthier S. Apolipoprotein E polymorphism and Alzheimer's disease. *Lancet*. 1993;342:697–9.
18. Roses AD. Alzheimer's disease: The genetics of risk. *Hosp Pract (Minneapolis)*. 1997;32:51–5, 58–63, 67–9.
19. Barabash A, Marcos A, Ancin I, Vazquez-Alvarez B, de Ugarte C, Gil P, et al. APOE, ACT and CHRNA7 genes in the conversion from amnesic mild cognitive impairment to Alzheimer's disease. *Neurobiol Aging*. 2009;30:1254–64.
20. Petersen RC, Smith GE, Ivnik RJ, Tangalos EG, Schaid DJ, Thibodeau SN, et al. Apolipoprotein E status as a predictor of the development of Alzheimer's disease in memory-impaired individuals. *JAMA*. 1995;273:1274–8.
21. Sasaki M, Kodama C, Hidaka S, Yamashita F, Kinoshita T, Nemoto K, et al. Prevalence of four subtypes of mild cognitive impairment and APOE in a Japanese community. *Int J Geriatr Psychiatry*. 2009;24:1119–26.
22. Tyas SL, Salazar JC, Snowdon DA, Desrosiers MF, Riley KP, Mendiondo MS, et al. Transitions to mild cognitive impairments, dementia, and death: findings from the Nun Study. *Am J Epidemiol*. 2007;165:1231–8.
23. Blom ES, Giedraitis V, Zetterberg H, Fukumoto H, Blennow K, Hyman BT, et al. Rapid progression from mild cognitive impairment to Alzheimer's disease in subjects with elevated levels of tau in cerebrospinal fluid and the APOE $\epsilon 4/\epsilon 4$ genotype. *Dement Geriatr Cogn Disord*. 2009;27:458–64.
24. Devanand DP, Pelton GH, Zamora D, Liu X, Tabert MH, Goodkind M, et al. Predictive utility of apolipoprotein E genotype for Alzheimer disease in outpatients with mild cognitive impairment. *Arch Neurol*. 2005;62:975–80.
25. Hamalainen A, Grau-Olivares M, Tervo S, Niskanen E, Pennanen C, Huuskonen J, et al. Apolipoprotein E $\epsilon 4$ allele is associated with increased atrophy in progressive mild cognitive impairment: a voxel-based morphometric study. *Neurodegener Dis*. 2008;5:186–9.
26. Hsiung GY, Sadovnick AD, Feldman H. Apolipoprotein E $\epsilon 4$ genotype as a risk factor for cognitive decline and dementia: data from the Canadian Study of Health and Aging. *CMAJ*. 2004;171:863–7.
27. Jack Jr CR, Petersen RC, Xu YC, O'Brien PC, Smith GE, Ivnik RJ, et al. Prediction of AD with MRI-based hippocampal volume in mild cognitive impairment. *Neurology*. 1999;52:1397–403.
28. Ramakers IH, Visser PJ, Aalten P, Bekers O, Sleegers K, van Broeckhoven CL, et al. The association between APOE genotype and memory dysfunction in subjects with mild cognitive impairment is related to age and Alzheimer pathology. *Dement Geriatr Cogn Disord*. 2008;26:101–8.
29. Tierney MC, Szalai JP, Snow WG, Fisher RH, Tsuda T, Chi H, et al. A prospective study of the clinical utility of APOE genotype in the prediction of outcome in patients with memory impairment. *Neurology*. 1996;46:149–54.
30. Slioter AJ, Cruts M, Kalmijn S, Hofman A, Breteler MM, Van Broeckhoven C, et al. Risk estimates of dementia by apolipoprotein E genotypes from a population-based incidence study: the Rotterdam Study. *Arch Neurol*. 1998;55:964–8.
31. Myers RH, Schaefer EJ, Wilson PW, D'Agostino R, Ordovas JM, Espino A, et al. Apolipoprotein E element 4 association with dementia in a population-based study: the Framingham Study. *Neurology*. 1996;46:673–7.
32. Rogaeva E, Meng Y, Lee JH, Gu Y, Kawarai T, Zou F, et al. The neuronal sortilin-related receptor sorl1 is genetically associated with Alzheimer disease. *Nat Genet*. 2007;39:168–77.
33. Harold D, Abraham R, Hollingworth P, Sims R, Gerrish A, Hamshere ML, et al. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nat Genet*. 2009;41:1088–93.
34. Lambert JC, Heath S, Even G, Campion D, Sleegers K, Hiltunen M, et al. Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. *Nat Genet*. 2009;41:1094–9.
35. Seshadri S, Fitzpatrick AL, Ikram MA, DeStefano AL, Gudnason V, Boada M, et al. Genome-wide analysis of genetic loci associated with Alzheimer disease. *JAMA*. 2010;303:1832–40.
36. Nuutinen T, Suuronen T, Kauppinen A, Salminen A. Clusterin: a forgotten player in Alzheimer's disease. *Brain Res Rev*. 2009;61:89–104.
37. Wollmer MA, Sleegers K, Ingelsson M, Zekanowski C, Brouwers N, Maruszak A, et al. Association study of cholesterol-related genes in Alzheimer's disease. *Neurogenetics*. 2007;8:179–88.
38. DeMattos RB, Cirrito JR, Parsadanian M, May PC, O'Dell MA, Taylor JW, et al. ApoE and clusterin cooperatively suppress A β levels and deposition: evidence that ApoE regulates extracellular A β metabolism in vivo. *Neuron*. 2004;41:193–202.
39. Pant S, Sharma M, Patel K, Caplan S, Carr CM, Grant BD. Amph-1/amphiphysin/Bin1 functions with RME-1/Ehd1 in endocytic recycling. *Nat Cell Biol*. 2009;11:1399–410.
40. Tebar F, Bohlander SK, Sorkin A. Clathrin assembly lymphoid myeloid leukemia (CALM) protein: localization in endocytic-coated pits, interactions with clathrin, and the impact of overexpression on clathrin-mediated traffic. *Mol Biol Cell*. 1999;10:2687–702.
41. Crehan H, Holton P, Wray S, Pocock J, Guerreiro R, Hardy J. Complement receptor 1 (CR1) and Alzheimer's disease. *Immunobiology*. 2012;217:244–50.

42. McGeer PL, Rogers J. Anti-inflammatory agents as a therapeutic approach to Alzheimer's disease. *Neurology*. 1992;42:447–9.
43. • Hollingworth P, Harold D, Sims R, Gerrish A, Lambert JC, Carrasquillo MM, et al. Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. *Nat Genet*. 2011;43:429–35. *This study implicated genetic variants in ABCA7, MS4A6A, EPHA1, CD33, and CD2AP in Alzheimer's disease.*
44. • Naj AC, Jun G, Beecham GW, Wang LS, Vardarajan BN, Buross J, et al. Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. *Nat Genet*. 2011;43:436–41. *This study implicated genetic variants in ABCA7, MS4A6A, EPHA1, CD33, and CD2AP in Alzheimer's disease.*
45. Cao H, Crocker PR. Evolution of CD33-related siglecs: regulating host immune functions and escaping pathogen exploitation? *Immunology*. 2011;132:18–26.
46. von Gunten S, Bochner BS. Basic and clinical immunology of siglecs. *Ann N Y Acad Sci*. 2008;1143:61–82.
47. Vitale C, Romagnani C, Falco M, Ponte M, Vitale M, Moretta A, et al. Engagement of p75/AIRM1 or CD33 inhibits the proliferation of normal or leukemic myeloid cells. *Proc Natl Acad Sci U S A*. 1999;96:15091–6.
48. Yamazaki T, Masuda J, Omori T, Usui R, Akiyama H, Maru Y. EphA1 interacts with integrin-linked kinase and regulates cell morphology and motility. *J Cell Sci*. 2009;122:243–55.
49. Adams RH, Klein R. Eph receptors and ephrin ligands. Essential mediators of vascular development. *Trends Cardiovasc Med*. 2000;10:183–8.
50. Duffy SL, Coulthard MG, Spanevello MD, Herath NI, Yeadon TM, McCarron JK, et al. Generation and characterization of EphA1 receptor tyrosine kinase reporter knockout mice. *Genesis*. 2008;46:553–61.
51. Holen HL, Nustad K, Aasheim HC. Activation of EphA receptors on CD4⁺CD45RO⁺ memory cells stimulates migration. *J Leukoc Biol*. 2010;87:1059–68.
52. Sakamoto A, Sugamoto Y, Tokunaga Y, Yoshimuta T, Hayashi K, Konno T, et al. Expression profiling of the ephrin (EFN) and Eph receptor (EPH) family of genes in atherosclerosis-related human cells. *J Int Med Res*. 2011;39:522–7.
53. Lehtonen S, Zhao F, Lehtonen E. CD2-associated protein directly interacts with the actin cytoskeleton. *Am J Physiol Renal Physiol*. 2002;283:F734–43.
54. Dustin ML, Olszowy MW, Holdorf AD, Li J, Bromley S, Desai N, et al. A novel adaptor protein orchestrates receptor patterning and cytoskeletal polarity in T-cell contacts. *Cell*. 1998;94:667–77.
55. Tanaka N, Abe-Dohmae S, Iwamoto N, Yokoyama S. Roles of ATP-binding cassette transporter A7 in cholesterol homeostasis and host defense system. *J Atheroscler Thromb*. 2011;18:274–81.
56. Chan SL, Kim WS, Kwok JB, Hill AF, Cappai R, Rye KA, et al. ATP-binding cassette transporter A7 regulates processing of amyloid precursor protein in vitro. *J Neurochem*. 2008;106:793–804.
57. Farrer LA, Cupples LA, Haines JL, Hyman B, Kukull WA, Mayeux R, et al. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. *JAMA*. 1997;278:1349–56.
58. Lee JH, Cheng R, Barral S, Reitz C, Medrano M, Lantigua R, et al. Identification of novel loci for Alzheimer disease and replication of CLU, PICALM, and BIN1 in Caribbean Hispanic individuals. *Arch Neurol*. 2011;68:320–8.
59. • Reitz C, Jun G, Naj A, Rajbhandary R, Vardarajan BN, Wang LS, et al. Variants in the ATP-binding cassette transporter (ABCA7), apolipoprotein E ϵ 4, and the risk of late-onset Alzheimer disease in African Americans. *JAMA*. 2013;309:1483–92. *This is the largest GWAS in African Americans implicating ABCA7 in Alzheimer's disease.*
60. Schurr E, Gros P. A common genetic fingerprint in leprosy and Crohn's disease? *N Engl J Med*. 2009;361:2666–8.
61. Capen EC, Clapp RV, Campbell WM. Competitive bidding in high-risk situations. *J Pet Technol*. 1971;23:641–53.
62. Xiao R, Boehnke M. Quantifying and correcting for the winner's curse in genetic association studies. *Genet Epidemiol*. 2009;33:453–62.
63. Zhong H, Prentice RL. Bias-reduced estimators and confidence intervals for odds ratios in genome-wide association studies. *Biostatistics*. 2008;9:621–34.
64. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 2007;447:661–78.
65. Hoggart CJ, Clark TG, De Iorio M, Whittaker JC, Balding DJ. Genome-wide significance for dense SNP and resequencing data. *Genet Epidemiol*. 2008;32:179–85.
66. Hong MG, Reynolds CA, Feldman AL, Kallin M, Lambert JC, Amouyel P, et al. Genome-wide and gene-based association implicates *FRMD6* in Alzheimer disease. *Hum Mutat*. 2012;33:521–9.
67. Lambert JC, Grenier-Boley B, Harold D, Zelenika D, Chouraki V, Kamatani Y, et al. Genome-wide haplotype association study identifies the *FRMD4A* gene as a risk locus for Alzheimer's disease. *Mol Psychiatr*. 2012;18:461–70.
68. Maher B. Personal genomes: the case of the missing heritability. *Nature*. 2008;456:18–21.
69. • Guerreiro R, Wojtas A, Bras J, Carrasquillo M, Rogaeva E, Majounie E, et al. *TREM2* variants in Alzheimer's disease. *N Engl J Med*. 2013;368:117–27. *This study implicated a rare variant in TREM2 in Alzheimer's disease.*
70. • Jonsson T, Stefansson H, Steinberg S, Jonsdottir I, Jonsson PV, Snaedal J, et al. Variant of *TREM2* associated with the risk of Alzheimer's disease. *N Engl J Med*. 2013;368:107–16. *This study implicated a rare variant in TREM2 in Alzheimer's disease.*
71. Lupton MK, Proitsi P, Danilidou M, Tsolaki M, Hamilton G, Wroe R, et al. Deep sequencing of the nicastrin gene in pooled DNA, the identification of genetic variants that affect risk of Alzheimer's disease. *PloS One*. 2011;6:e17298.
72. Bettens K, Brouwers N, Engelborghs S, Lambert JC, Rogaeva E, Vandenbergh R, et al. Both common variations and rare non-synonymous substitutions and small insertion/deletions in CLU are associated with increased Alzheimer risk. *Mol Neurodegener*. 2012;7:3.